

# Automated Medical Waste Detection Using a YOLO-Based Machine Learning Approach

Dafni Panopoulou, Dionysis Xenakis, Dimitrios Panopoulos, Sotirios Kousouris

**Abstract**—The management of medical waste is a critical matter globally, as existing manual sorting techniques expose personnel to significant hazards, and may jeopardize community and environmental safety. To mitigate these risks, automated solutions are being developed, however progress seems to have been restrained by the lack of publicly available, annotated datasets suitable for object detection. This prohibits direct comparison of the models, and the development of more efficient ones. This paper directly addresses this challenge by developing and presenting a methodology for automated medical waste detection. Dataset A from public dataset “Medical Waste 4.0” was used to create a new annotated and augmented version suitable for object detection. Using this dataset, a high performance was established using the state-of-the-art YOLOv9c model achieving a mean Average Precision of 0.984. The impact of this work is twofold. It provides a highly accurate and efficient solution for an urgent safety matter, and also allows for comparison with other models, paving the way for advanced automated medical waste detection systems.

**Index Terms**—medical waste management, object detection, computer vision, machine learning, YOLOv9

## I. INTRODUCTION

The effective management and disposal of medical waste is a critical issue globally, as it can substantially affect sectors like public health, environmental safety, as well as personnel safety. It is mandatory for hazardous materials to be separated from general waste in healthcare facilities worldwide. This process usually relies on manual labor, rendering it time-consuming and cost-inefficient, but also dangerous, as it exposes health workers to significant risks, such as injuries from sharp items like needles and blades. Personnel could also be exposed to infectious pathogens or experience chemical burns. Another factor that should be considered is that human error in manual segregation of the items contributes to improper waste disposal, potentially leading to contamination of soil and water, thus posing a long-term threat to the environment and the community. According to global health organizations, it is essential to develop safe, effective, and reliable waste segregation for the healthcare sector. It is crucial to ensure the proper management and disposal of medical waste. An additional amount of 87000 tonnes of personal protective equipment (PPE) waste was generated globally with the COVID-19 pandemic [1]. This matter affects not only public health, but also environmental safety and occupational security to a great extent. It is essential that hazardous materials, which account for approximately 15% of the total medical waste, are segregated from general medical waste in healthcare facilities worldwide [2]. The process required to achieve this is both logistically complex and cost-consuming. The global medical waste management market, which comprises all services from collection to disposal of

medical waste, was valued at approximately 16.3 billion United States Dollars (USD) in 2023, and is estimated to reach 27.8 billion by 2032 with a Compound Annual Growth Rate (CAGR) of 6.1% [3]. The significant cost of the market can be attributed to a great extent to the fact that the segregation process was executed manually, which is time-consuming and exposes the personnel to severe risks. A significant risk associated is this is injuries from sharp objects such as needles and blades, with a study reporting that over 70% of waste handling personnel have experienced at least one sharp-related injury through their professional careers [4]. The inherent danger of the process necessitates the waste management process to be specialized. The handling and treatment of medical waste are the main contributors to the cost of the procedure, with treatment and disposal services amounting to 75% of the total market revenue [5]. Costs seem to range between 0.5 to 1.5 USD per pound for large-scale disposal. These numbers can be affected by region and treatment method [6]. Manual waste segregation is susceptible to human error, a factor that contributes to improper disposal, thus posing severe risks to the environment and public health. It is estimated that a significant part of this waste, which in an important amount is plastic-based, is managed and burned in an inappropriate manner. Plastic waste seems to amount for up to 15% to 25% of the total medical waste, with millions of tonnes of medical waste being generated every year. The incineration of one tonne of plastic medical waste may generate between 2.5 to 3 tonnes of CO<sub>2</sub> emissions, while alternative methods could potentially lead to water and soil contamination. Globally, medical waste management is handled with varying levels of infrastructure depending on the region. North America has stringent regulations regarding proper medical waste management and currently holds over 36% of the revenue share from services such as medical waste disposal and treatment. The Asia-Pacific region seems to have the biggest growth rate, which can be attributed to increasing global awareness and investment in proper medical waste disposal and treatment technologies [3]. In conclusion, the development of reliable, safe, and efficient medical waste management techniques is an urgent and important matter in order to protect human health, the community, and ultimately our planet.

### A. Problem Statement

Recent advancements in technology pave the way for the automation of the medical waste management process. Machine learning and computer vision techniques particularly show immense promise in automating complex visual tasks. While

machine learning has successfully been applied to image classification tasks, progress in this domain is hindered due to a specific issue. These methods are insufficient for robotic automation, as this requires precise object localization. There is advanced research on object detection, which provides this localization, however prior studies have almost exclusively created and utilized small-scale and private datasets. This prohibits direct model comparison, as the results cannot be reproduced, thus a barrier is created regarding collaborative scientific advancement.

Creating an automated system with a high performance in object detection can be a challenge, so two contributions are aimed to be delivered with this work. The main goal is to devise and implement a complete and effective methodology for medical waste detection, from data preparation stage to deploying the model. The contributions are:

- 1) A new, large-scale, and annotated version derived from the Medical Waste 4.0 dataset [7]. Bounding box annotations were added to 5516 images on dataset A of the aforementioned dataset, thus creating a new dataset suitable for object detection tasks.
- 2) A strong baseline. The YOLOv9c model was used, creating a baseline with very high performance. With its performance metrics, as will be analyzed in the results section, it seems to be a very effective solution.

The structure of this work is as follows. After the introduction, there is a review regarding the state of the art in object detection and medical management in section 2. The system model and the dataset is described in section 3. In section 4, the proposed solution is presented based on the YOLOv9c model. A detailed analysis of the results after the evaluation is presented in section 5. Finally, section 6, where conclusions and suggestions regarding future directions are presented.

## II. STATE OF THE ART

In the past few years, research in the field of medical waste management has progressed with the application of artificial intelligence (AI) techniques, especially computer vision and machine learning, as there is a growing need to enhance safety, reduce costs, and improve waste segregation efficiency. Various methodologies have been studied exploring possibilities that would allow for these advancements. A review of the existing literature regarding our field of interest shows that these studies can be divided into two main categories. Studies that used image classification techniques and studies that used object detection techniques. Both techniques have advantages and limitations that shape the current research landscape [8].

### A. Image classification for Medical Waste Management

Convolutional Neural Networks (CNNs) have been greatly used for image classification tasks regarding medical waste. In image classification, the main purpose is to assign a single and correct label to an image, in this case a medical waste item. This approach has shown promise, with models reaching very high metrics. A study used ResNeXt, which is a type of CNN, combined with transfer learning for results improvement. The task was to automatically classify different

types of medical waste using deep learning to achieve proper waste management. The process pipeline is the following. The images were collected, cleaned, and augmented. To continue, transfer learning was used with ResNeXt to train the model. Cross-validation was used for the evaluation, as well as other metrics. Finally, the trained model was deployed to classify medical waste images automatically. The dataset is private and contains 3480 images collected from the First Affiliated Hospital, Zhejiang University in 2019. It consists of eight classes, which are types of medical waste. The datasets contain some images where more than one item is in the same image, which could affect accuracy. However, the model achieved an exceptional 97.2% accuracy in classifying medical waste, thus contributing to the field. However, the dataset size is moderate, and some images have poor quality, which allows for further changes and improvement [9]. In another study, the task was to develop a system for medical waste classification. The dataset was private and custom-built due to the absence of publicly available datasets for medical waste. It contains eleven images and amounts to 3200 images collected from Google Images, which were manually sorted to ensure quality. Data augmentation techniques were applied, such as affine transformations and color modifications to increase dataset diversity and prevent overfitting. Several pre-trained CNN models were employed for the training process, including EfficientNetB7, InceptionV3, ResNet50, and VGG19. Among these, EfficientNetB7, InceptionV3, and ResNet50 achieved accuracy levels above 95%. The process pipeline is the following. The images were collected using JavaScript and Python scripts, with manual sorting to ensure quality. Then the data were augmented to expand the size and diversity of the data. To continue, pre-trained models were used, and finally, the system was deployed for medical waste segregation, allowing images to be uploaded and classified instantly. A study developed an IoT solution, however it focuses on image classification, which is not as practical in real-world settings [10]. Furthermore, the first paper that created and utilized the Medical Waste 4.0 Dataset, which is the source of the dataset used in this work, employed a deep learning CNN based on EfficientNet and achieved an accuracy of 100% on an image classification task. The task was to classify medical waste items to help operators with waste segregation. The images were collected using a special stereo camera setup. The process pipeline was the following. The images were collected, then prepared and resized for the model, the model was trained and then used to classify items in real-time. The results achieved are impressive and show that medical waste items have distinct characteristics and visual features. A significant contribution from the study is the creation of a medical waste dataset that can be used for object classification. However, an image classification does not provide any information regarding the location of an object. As noted by the authors of the aforementioned study this is a significant drawback for real-world scenarios, as for an automated waste segregation and disposal system to work, the robot needs to know exactly where the object is, not only what kind of object it is. This limitation

emphasizes that future directions should be towards object detection.

### *B. Object Detection for Medical Waste Management*

A number of studies have focused on object detection as it is considered an effective way to manage medical waste. Object Detection allows the classification and the localization of the object using bounding boxes, which is more suitable for developing waste segregation automated systems. A variety of architectures and models have been researched regarding their capabilities in these studies. In a study OpenCV was used to detect objects in images. The task was to detect and classify medical waste items using machine vision. The objects were located with morphological filtering, and then classified using the SSD-MobileNet. The dataset was approximately 2800 images, including four types of medical waste. The images were also augmented with rotations to enhance diversity. The model achieved over 98.5% accuracy. However, the dataset is private and limited to certain types of medical waste and may not be practical for real-world scenarios [11]. In another study, the task was to build an autonomous system that is capable of identifying and separating medical waste into four types, which are: infectious, hazardous, radioactive, and general waste. The dataset is private and custom-made, collected from Google Images, COCO dataset, and self-collected images. After the collection, the images were labeled using labelling into four types of medical waste. The images were then resized and augmented. Then, machine learning classifiers were used, like K-Nearest Neighbors (KNN), Naive-Bayes, to classify the images into the four categories. Furthermore, the system was integrated with a robot that included a camera and a robotic arm to detect waste and segregate it accordingly. Regarding the detection LiDAR technology was used. This study contributes by focusing on the hardware implementation, however the AI component is a classifier and the dataset is small and private [12]. A prominent trend recently within the field of medical waste detection has been the development of lightweight efficient models, which will be suitable and effective regarding deployment embedded systems. One of the aforementioned works, is a study, in which a COVID-19 waste detection system is developed using YOLOv5 on Raspberry Pi. The dataset is private and includes 7 types of waste and initially amounted to 175 images. After the augmentations it increased to 1050 images. The process pipeline is as follows: The images were collected, annotated using makesense.ai, augmented by rotating images, scaling them, and adding noise. The augmented data was the input for the YOLOv5 model, which after training was converted for Raspberry Pi. A real-time detection was executed with a USB on the Raspberry Pi. Finally, the model was evaluated and achieved an average accuracy of over 96%. This study showed that YOLOv5 can be executed effectively on Raspberry Pi for waste classification, however the dataset is small and private, and the accuracy scores were lower when the images were smaller (85%) [13]. In another study, the YOLOv4-tiny model was used in order to develop a lightweight algorithm for embedded applications, with an achieved mean Average Precision (mAP) of 93.9% [8].

The task was to detect and classify medical waste objects in images. The dataset is private and contains 1500 images of six different medical waste categories. The process pipeline is as follows: The data was collected and annotated, the YOLOv4-tiny was modified and trained using transfer learning with pre-trained weights. The performance of the model was then evaluated, reaching a 93.9% detection accuracy. The model was then deployed and implemented for real-time medical waste sorting. A computationally efficient model for embedded systems was created, however the dataset was private, limited in size and the model that was model is older so a higher accuracy may have been possible [14]. Furthermore, a study utilized the YOLOv7 model for object detection and classification tasks with video streams with medical waste items falling which were collected in a hospital setting to create the "iWaste" system. The dataset contains 970 videos with duration of five seconds at 24 frames per second, of medical waste falling into a bin. They were captured with a stationary camera in the operating room (OR). The items may belong to one of four classes. After the videos were collected, motion detection was used to detect frames of the object falling. Relevant frames were extracted and trimmed to focus on the object. Then the hybrid neural network was trained R3D+C2D, where C2D is a 2D convolutional network and R3D a version of ResNet adapted to handle 3D convolutions, and achieved an accuracy of 79.99%. The dataset is private, small and contains only four types, however, this work is the first to use video-based detection of medical waste in OR settings [15]. In another study a complete system was developed, called Smart Waste Manager. The task was to detect and classify medical waste images. The dataset was 2000 images and consisted of 8 classes. The data was collected, labeled and augmented. The YOLOv8 model was utilized and used for real-time waste detection and achieved an accuracy of over 97.14%. The dataset is small and private, however, this work shows potential for application in real-world scenarios [16]. In another study the task was to classify medical waste using images on an IoT-based system. The dataset was custom-made and consisted of 1937 images collected from Google and Bing, and separated in four classes. The images were resized and normalized for the training. Five pre-trained models were trained, including ResNeXt, EfficientNet, ConvNext, ResNet and ViT b 16. ResNeXt achieved a slightly higher predictive accuracy. ResNeXt was trained on the dataset achieving an accuracy of 97.93%. The ResNeXt was then integrated in the SmartMedWaste system, which includes IoT components for automated medical waste management. The process pipeline is as follows: The sensors trigger the camera when they detect medical waste. The camera captures the image, in which the item is detected. The model classifies the type of medical waste, and the robot segregates it into proper bins. The dataset is custom-made and may not reflect real conditions, however, a system that combines sensors, AI, and robotics has been successfully created [17]. Research in the field has greatly advanced in the last few years regarding medical waste detection with the utilization of image classification and object detection techniques. However, a challenge remains that

potentially hinders faster and greater advances in the field. Most studies utilized private, custom-collected datasets instead of standard ones that would be accessible for reproducibility purposes. Even though this allows for refined developed solutions, the research landscape is limited, as it is impossible to compare the performance of different models, reproduce results with the same model, or elevate prior work when utilizing a different dataset. This study contributes to the field by developing and presenting a methodology for creating a large-scale annotated dataset, and using it to establish a new performance benchmark utilizing a YOLO model. The 5516 images of dataset A from the Medical Waste 4.0 dataset were annotated using both automated and manual techniques. To continue, with the exceptional results achieved by the highly efficient YOLOv9c model, a solution achieving a high performance and a modern point of comparison is provided which contributes in the research of automated medical waste detection.

### III. MEDICAL WASTE DETECTION PIPELINE

#### A. Data Acquisition and Object Classes

The initial dataset is the "Medical Waste 4.0 Dataset", which is publicly available and was created by a group of researchers at CNR-ISTI, Italy, under the framework of a project funded by the Tuscany region [7], [18]. The dataset was created in order to help develop new AI methods and optimize the already existing ones to train, validate and test machine learning algorithms for medical waste classification. Regarding the equipment that was used, the images were taken using an OAK-D stereo camera with depth perception capabilities, however, only the standard RGB images were used in this work. The initial dataset consists of two smaller datasets, Dataset A and Dataset B. Dataset A is the main dataset that can be used to train, validate and test machine learning algorithms for medical waste classification. Dataset B was acquired on different days from Dataset A, in order to avoid any possible duplicated images between datasets A and B. It can be used for additional final testing. This research utilizes Dataset A. There are 5516 images in Dataset A, and in each of them only one medical waste item is portrayed against a uniform white to light grayish background. This controlled environment allows for the minimization of background noise and a more focused evaluation of the model's ability to distinguish and detect each class. The Dataset consists of 13 classes, including a range of items which are common in a medical environment. The classes are the following: Gauze, Glove pair latex, Glove pair nitrile, Glove pair surgery, Glove single latex, Glove single nitrile, Glove single surgery, Medical cap, Medical glasses, Shoe cover pair, Shoe cover single, Test tube, Urine bag.

#### B. Annotation Methodology

The images in Dataset A were already classified so it was essential to add the bounding boxes in the images in order to make the dataset suitable for object detection. Due to the number of images the process that was followed was semi-automated. The images were uploaded in Roboflow website to be annotated. Without the annotations, the images would only be suitable for image classification, not for object detection.

Initially, the images were labeled automatically with Grounding DINO model. Each class was uploaded separately and a small simple description of the object was the input to the Grounding DINO model in order to locate this described object in the picture and label it. To continue, the images and their proposed annotations were further investigated and thoroughly inspected to ensure the annotations were correct. Furthermore, there were images where the model failed to add the annotation even though there was an object in the picture. In these cases a bounding box was drawn manually through roboflow [19]. Creating this dataset and detailing the methodology that was used for this purpose, allows for comparison with other models, to examine and evaluate their performance. Most studies examined used private, custom-made datasets that did not allow for further comparison and evaluation between models [11], [12], [13], [14], [15], [16].

#### C. Data Segmentation

In machine learning, the data segmentation process is a fundamental step in order to properly evaluate a model. The assessment regarding the model's ability to generalize on new, unseen data needs to be unbiased, instead of evaluating how well it memorized the data it was trained on. The evaluation needs a new set of unseen data, otherwise the performance metrics would be inflated and would not be objective. The standard technique is to separate the dataset into three subsets. These would be the training set, the evaluation set, and the test set. The training set usually amounts to 70-80% of the data, and is used to train the model.

The validation set usually amounts to 10-20% of the dataset and is used during the training process to adjust the model's parameters and prevent overfitting. The test set usually amounts to 10-20% of the dataset and is used only once, after the training is completed, or it could be unseen data unrelated to the initial dataset. It is used to evaluate the final performance of the model. The process directly affects the validity and objectivity of the final performance results. The test set must never be used during the training process, and any other process before the final evaluation, to ensure the model has never seen these data before, so that the final performance metrics are representative of the model's performance on newdata. A review of the relevant literature in the medical waste detection domain confirms that data segmentation is indeed a fundamental step in developing machine learning (ML) models. The standard practice is to separate the dataset into at least a training set and a validation set. The test set can either belong to the initial dataset or may be completely unrelated to it. Various splits were used for training and testing purposes, respectively, e.g. 80% of the dataset for training and 20% for testing [13], 88%/12% [15]. In the relevant literature, the dataset would be separated in three parts: the training set, the validation set, and the test set. This study on a smart waste manager system executed a 60/10/30 split for training, evaluation, and testing, respectively [17]. Nearly all reviewed papers imply a data segmentation, however, the precise ratios that were used were not mentioned in some [11], [12], [16], [14]. Clearly defining the split, as

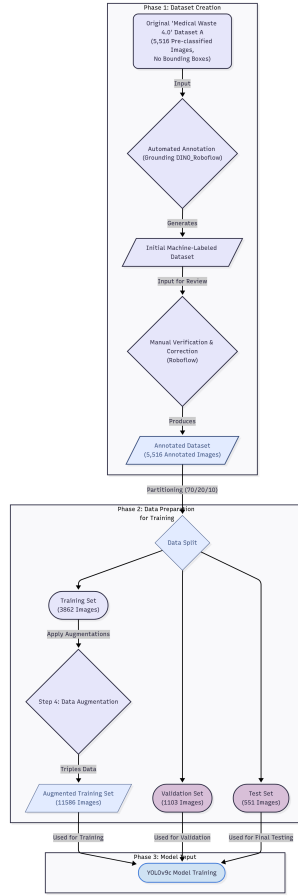


Fig. 1. Medical Waste Detection Pipeline

in this work, is aligned with the standard methodology in the field. The data segmentation and augmentation strategy that was executed in this work was designed to develop a highly accurate and efficient model. The dataset was initially separated in three parts, the training set, the validation set, and the test set, 70/20/10 respectively. This ensures that the majority of the images is available for the model to be trained on and learn from. Furthermore, the training set tripled in size through the data augmentation process, a process that enhances data diversity in order to help the model to generalize better and prevent overfitting. To continue, using a validation set that amounts to 20% of the dataset provides a stable and reliable way to evaluate the model's performance during training. The test set amounts to 10% of the dataset, and its size ensures a reliable and unbiased final evaluation.

#### IV. PROPOSED SOLUTION

As the automation of medical waste detection is a prominent challenge, this work emphasizes a solution based on machine learning techniques. It proposes a solution using a state-of-the-art detection model from the YOLO family, specifically, YOLOv9c, because of its performance and stability. In this section, there will be further analysis on the model architecture, the data augmentation strategies that were employed to enhance the model's performance, how the training was executed, and how the training was evaluated.

#### A. YOLOv9c Architectural Overview

The YOLOv9c model is one of the latest advancements in the YOLO family. Regarding its architecture, it can be divided into three main components, which would be the backbone, the neck, and the head. For this work, YOLOv9c was used, as it seems to exhibit an optimal balance between size and performance. This variant has approximately 25.3 parameters, and a computational cost of 102.4 GFLOPs [20]. The basic components of its architecture, as aforementioned, will be analyzed further below. The image, before it is utilized as input, is resized to a fixed size and normalized, which ensures uniformity and faster processing. Each image is then divided into smaller cells, and each cell is responsible for detecting objects within its boundaries.

**Backbone:** The backbone is the architecture that handles feature extraction from the input images.

**Auxiliary:** The auxiliary architecture is an additional one that is added to the YOLOv9 version of the YOLO family. It allows it to enhance the reliability of the training process by additional information that links the information of the input images to the output target.

**Neck:** The neck architecture is built by utilizing a pyramid network method, which is used to combine features from multiple layers of the backbone model.

**Head:** The head architecture is responsible for performing the detection of objects based on the features from the neck. It generates the output, the final vectors that define the bounding boxes of the detected objects. It also generates the confidence score for each bounding box, as well as the classification probabilities for each of the 13 classes that exist in this work.

#### B. Data Augmentation Strategy

Data augmentations were done to the dataset to improve the model's performance and generalization capability and prevent overfitting. The dataset tripled in size, as for each image, three outputs were created. This increase regarding the diversity of the data allowed the model to learn more invariant and robust features. The augmentations that were done to the data were mostly geometric and photometric and will be further discussed below:

**Geometric Data Augmentations:** The data were altered geometrically using two techniques. These techniques were crop and rotation. The images were randomly cropped with the zoom factor ranging between 0%, which was no zoom at all, up to 20% zoom. The images were randomly rotated with the rotation factor ranging between -15 up to +15 degrees.

**Photometric Data Augmentations:** The images were altered photometrically using two techniques, to change the visual properties of the image and simulate different lighting conditions. The visual properties that were altered were brightness and exposure. The brightness of the images was randomly adjusted, with the brightness factor ranging between -15% +15%. The exposure of the images was randomly adjusted, with the exposure factor ranging from -10% up to +10%.

**Mosaic and MixUp Data Augmentation:** YOLOv9 employs techniques like mosaic and mixup augmentation to

improve model robustness, reduce overfitting and enhance its generalization to real-world scenarios. Mosaic augmentation allows four images to be combined into one, which allows the model to learn how to detect the object in cluttered and partially occluded spaces. MixUp augmentation creates a composite image that is the result of two other images blending.

### C. Training

The environment that was chosen for training was Google Colab, as it is a cloud computing environment and allows utilization of NVIDIA A100 GPU, which will reduce the time needed for the model to be trained. The Ultralytics framework was used to train the YOLOv9c model, and PyTorch. To prevent overfitting and reduce unnecessary training time, the model has an early stopping mechanism. Patience parameter enables the model to stop training if no improvement is observed regarding the primary evaluation metric after a chosen number of consecutive epochs, defined as patience. In this work, the number of epochs was set to 300 and patience was set to 50. The model's weights are updated, by an optimization algorithm, based on the computed loss. To ensure stable convergence, the step size of the aforementioned updates is adjusted by a learning rate scheduler. The values for the optimizer and the learning rate were the values that are provided by the Ultralytics framework by default.

### D. Loss Function

In order for the training process to be most advantageous, the total loss during the training process needs to be minimized. The total loss consists of three primary components which will be analyzed further below. The output logs from the training process after each epoch provide direct insights regarding these metrics, allowing the model to be optimized. Based on these outputs, the optimizer changes the model's parameters, also known as its weights. These weights are adjusted so that the losses are minimized.

**Bounding Box Regression Loss (box loss):** This loss measures the difference between the predicted bounding box and the ground-truth box. The differences can be any inaccuracies regarding the box's position and its coordinates, its size, and aspect ratio.

**Classification Loss (cls loss):** This loss measures whether the object was classified correctly within a given bounding box. If an object is correctly detected but incorrectly classified, it can be measured by this type of loss.

**Distribution Focal Loss (dfl loss):** Box coordinates are usually presented as single values. However, in order to capture the inherent uncertainty in localization, the model learns a probability distribution around each coordinate instead of single values. Distribution focal loss assigns more weight in rare and challenging instances, helping the model to focus the training and eventually leading to more accurate, and confident object localization even with challenging instances. The sum of the above three losses is then calculated with the aim to minimize it, which is what the optimizer does, in order to improve the model's ability to find objects, classify them correctly, and place accurate bounding boxes around them.

## V. EXPERIMENTS AND RESULTS

The model was evaluated, and the results are presented in this section. The evaluation was conducted using the test set, that was held out and not used in the training process, meaning that these were new images that the model had not seen before, to ensure that the assessment of the model's performance was unbiased.

### A. Experimental Setup

The evaluation, like the training process, was conducted in a Google Colab environment, utilizing an NVIDIA A100 GPU, and also PyTorch and Ultralytics frameworks. The model that was evaluated is a trained YOLOv9c, as explained previously, with 156 layers, approximately 25 million parameters, and a computational complexity of 102.4 GFLOPs.

### B. Overall Quantitative Performance

The performance of the model when evaluated using the test set was exceptionally high, which indicates its effectiveness regarding medical waste detection. Performance metrics of the model will be further analyzed below. The most important metric is considered to be mean average precision  $mAP@0.50:0.95$  where the model scored 0.984, indicating very good accuracy regarding the classification and the localization of objects. The  $mAP@0.50$  metric is more lenient as it only requires a 50% overlap threshold, so the model scored higher reaching 0.994. Furthermore, the average recall reached by the model was 0.994, which indicates that it correctly identified most of the objects that were to be detected. In addition, the inference speed of the model is 4.5 ms/image, which potentially makes it suitable for real-time applications.

### C. In-Depth Class-wise Error Analysis (confusion matrix)

For a more detailed analysis a diagram was created to visualize the results after the evaluation on the test set. A confusion matrix presented in Fig.2, was used, in order to have a better understanding of the model's behavior and how it performed with each class. The high accuracy, across all 13 classes, when measuring the classification ability of the model can be observed from the diagonal, where all values are above 0.98. For some classes the model achieved a recall score of 1. This means that the model was able to identify correctly every single instance of these classes in the test set. These classes were: medical cap, glove pair surgery, and glove single surgery. Regarding misclassifications, more information on those can be obtained from the off-diagonal elements. Some classes are visually similar and this causes confusion, hence errors. For example, two instances of the class glove single latex seem to have been misclassified as gauze. This is a logical error, as the two classes share some characteristics regarding their color properties or perhaps even texture. It is also quite logical for some misclassifications to exist between single objects and their paired counterparts, especially if only one object is clearly visible in the picture. Furthermore, from the analysis we can get some insights regarding errors related to the background. The number of false positives generated from the model was small. The most frequent ones were four instances where glove single nitrile instances were missed and classified as

background. However, the number of false negatives generated by the model shows that only in very few cases did the model miss an object entirely. To conclude, based on the confusion matrix, the model seems to be highly accurate with errors that can occur in logical scenarios, meaning visually ambiguous or similar objects.

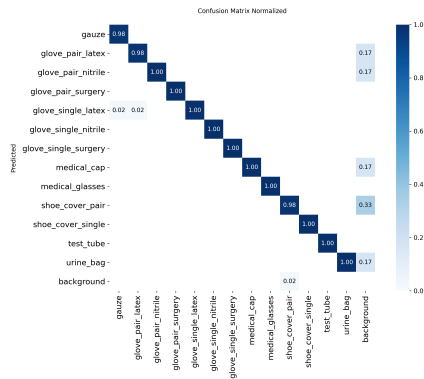


Fig. 2. Confusion matrix

## VI. CONCLUSIONS AND FUTURE DIRECTIONS

This work addressed the need for an efficient solution regarding medical waste management, specifically, medical waste detection using AI techniques. A complete methodology was presented, from the acquisition, annotation and adaptation of a large-scale dataset, and continuing with the deployment of the YOLOv9c model. The model was evaluated and achieved an exceptional mAP of 0.984 on the dataset. The inference speed was only 4.5 milliseconds per image which means that the model is not only accurate, but also fast and efficient enough potentially for real-world applications, making this work a strong foundation for future advancements in this domain.

Future directions could include, but are not limited to:

- **Model comparison:** The performance of other models on the same dataset should be tested and evaluated to properly contextualize the model's performance. Highly efficient CNN-based models like EfficientDet, and Transformer-based models like DETR should be tested in order to determine whether their differences yield a performance advantage on this dataset or not.
- **Dataset Expansion:** The performance of any model is highly dependent on the input it receives. This includes the characteristics and the size of the training dataset. In future works, the dataset could be further expanded to include more complex scenarios, such as a cluttered background, which would simulate real-world environments more accurately.
- **Real-World Deployment and Video Analysis:** Future research could also include model deployment on a real-time sorting system, potentially on edge computing devices. Real-time object tracking could be realized in video streams of sorting systems.

## ACKNOWLEDGEMENT

This work received funding from Horizon Europe under the Marie Skłodowska Curie actions: ELIXIRION (GA 101120135).

## REFERENCES

- [1] "Global analysis of health care waste in the context of COVID-19." Accessed: Jul. 10, 2025. [Online]. Available: <https://www.who.int/publications/i/item/9789240039612>.
- [2] E. Janik-Karpinska et al., "Healthcare Waste—A Serious Problem for Global Health," *Healthcare*, vol. 11, no. 2, p. 242, Jan. 2023, doi: 10.3390/healthcare11020242.

- [3] "Medical Waste Management Market to Hit to USD 26.08 Bn by 2034." Accessed: Jul. 10, 2025. [Online]. Available: <https://www.precedenceresearch.com/medical-waste-management-market>.
- [4] "(PDF) Occupational Exposure to Needlestick and Sharp Injuries among Hospital Waste Handlers in Selected Government Health Facilities of Bhopal District," ResearchGate. Accessed: Jul. 10, 2025. [Online]. Available: [https://www.researchgate.net/publication/344215138\\_Occupational\\_Exposure\\_to\\_Needlestick\\_and\\_Sharp\\_Injuries\\_among\\_Hospital\\_Waste\\_Handlers\\_in\\_Selected\\_Government\\_Health\\_Facilities\\_of\\_Bhopal\\_District](https://www.researchgate.net/publication/344215138_Occupational_Exposure_to_Needlestick_and_Sharp_Injuries_among_Hospital_Waste_Handlers_in_Selected_Government_Health_Facilities_of_Bhopal_District).
- [5] "Medical Waste Management Market to Hit to USD 26.08 Bn by 2034." Accessed: Jul. 10, 2025. [Online]. Available: <https://www.precedenceresearch.com/medical-waste-management-market>.
- [6] "Stericycle." Accessed: Jul. 10, 2025. [Online]. Available: <https://www.stericycle.com/en-us>.
- [7] X. Bian, Y. Chen, S. Wang, F. Cheng, and H. Cao, "Medical Waste Classification System Based on OpenCV and SSD- MobileNet for 5G," in 2021 IEEE Wireless Communications and Networking Conference Workshops (WCNCW), Mar. 2021, pp. 1–6. doi: 10.1109/WCNCW49093.2021.9420036.
- [8] S. Bobe, P. Adhav, O. Bhalerao, and S. Chaware, "Review on Deep Learning Based medical Waste Detection and Classification," presented at the Proceedings of the 2nd International Conference on Edge Computing and Applications, ICECAA 2023, 2023, pp. 1071–1076. doi:10.1109/ICECAA58104.2023.10212343.
- [9] H. Zhou et al., "A deep learning approach for medical waste classification," *Sci. Rep.*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-06146-2.
- [10] M. Verma, A. Kumar, and S. Kumar, "Medical Waste Classification using Deep Learning and Convolutional Neural Networks," in 2022 IEEE Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), Dec. 2022, pp. 1–5. doi: 10.1109/IATMSI56455.2022.10119431.
- [11] X. Bian, Y. Chen, S. Wang, F. Cheng, and H. Cao, "Medical Waste Classification System Based on OpenCV and SSD- MobileNet for 5G," in 2021 IEEE Wireless Communications and Networking Conference Workshops (WCNCW), Mar. 2021, pp. 1–6. doi: 10.1109/WCNCW49093.2021.9420036.
- [12] S. S. C S, S. C S, S. S, and U. S, "Automatic Bio Medical Waste Segregator," in 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Mar. 2022, pp. 1124–1130. doi: 10.1109/ICACCS54159.2022.9785342.
- [13] I. Hermawan, A. Mardiyono, R. W. Iswara, F. A. Murad, M. A. Ardiawan, and R. Puspita, "Development of Covid Medical Waste Object Classification System Using YOLOv5 on Raspberry Pi," presented at the 2023 10th International Conference on Information Technology, Computer, and Electrical Engineering, ICITACEE 2023, 2023, pp. 443–447. doi: 10.1109/ICITACEE58587.2023.10277207.
- [14] Y. S. Gan, Y.-Z. Liu, B.-C. Tseng, G.-B. Liong, and S.-T. Liong, "Lightweight Deep Learning Algorithm for Sorting Medical Waste Embedded System," *J. Internet Technol.*, vol. 25, no. 5, pp. 683–700, 2024, doi: 10.70003/160792642024092505004.
- [15] J. Chen, J. Mao, C. Thiel, and Y. Wang, "iWaste: Video-Based Medical Waste Detection and Classification," in 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & logy Society (EMBC), Jul. 2020, pp. 5794–5797. doi: 10.1109/EMBC44109.2020.9175645.
- [16] S. Bobe, P. Adhav, O. Bhalerao, and S. Chaware, "Smart Waste Manager: YOLO-driven Medical Waste Detection and Classification System," in 2024 IEEE International Conference on Smart Power Control and Renewable Energy (ICSPCRE), Jul. 2024, pp. 1–6. doi: 10.1109/ICSPCRE62303.2024.10674956.
- [17] S. K. Karmakar, M. A. Mallick, A. I. Middy, and S. Roy, "AI Powered Medical Waste Management with IoT," in 2024 IEEE Calcutta Conference (CALCON), Dec. 2024, pp. 1–6. doi: 10.1109/CALCON63337.2024.10914143.
- [18] A. Bruno, C. Caudai, G. R. Leone, M. Martinelli, D. Moroni, and F. Crotti, "Medical Waste Sorting: A Computer Vision Approach For Assisted Primary Sorting," in 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), Jun. 2023, pp. 1–5. doi: 10.1109/ICASSPW59220.2023.10193520.
- [19] B. Dwyer, J. Nelson, and T. Hansen, Roboflow. (2025) Roboflow. [Online]. Available: <https://roboflow.com>.
- [20] "What is YOLOv9: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector." Accessed: Jul. 07, 2025. [Online]. Available: <https://arxiv.org/html/2409.07813v1>.